DOCUMENT RESUME

ED 393 940 TM 024 982

AUTHOR Rakow, Ernest A.
TITLE Unequal Cell Size ANOVA and the Meaning of the Interaction.
PUB DATE 9 Nov 95
NOTE 17p.; Paper presented at the Annual Meeting of the Mid-South Educational Research Association (24th, Biloxi, MS, November 1995).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Analysis of Variance; *Estimation (Mathematics); *Interaction; *Research Design
IDENTIFIERS Statistical Analysis System; Statistical Package for the Social Sciences; *Unequal Cell Size; Weighting (Statistical)

ABSTRACT

Analysis of variance (ANOVA) is a frequently used statistical procedure in education and the social sciences. Very often the use of ANOVA involves situations with unequal cell sizes. When confronted with data to analyze from an unbalanced design, the researcher should select very carefully from the method or option in the statistical package being used for estimation of the sums of squares. As is illustrated, conflicting results may occur depending on the options chosen in the analysis. One approach for doing ANOVA is the use of unweighted means analysis. Interaction effects in such analysis are explored. Authors who have most clearly explained the alternatives for analyzing designs with unequal cell sizes have done so by emphasizing comparisons of models through the use of sums of squared errors for competing models, and they have also tended to use regression methodology to explicate the various models for nonorthogonal designs. Some of their points are reviewed through examples using the Statistical Analysis System and the Statistical Package for the Social Sciences. (Contains 8 tables and 13 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

UNEQUAL CELL SIZE ANOVA AND THE MEANING OF THE INTERACTION

ED 393 940

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

ERNEST A. RAKOW

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC).

Ernest A. Rakow

The University of Memphis

Paper presented at Twenty-Fourth Annual Meeting of the Mid-South Educational Research Association, Biloxi, Mississippi, November 9, 1995.

TM 024982

BEST COPY AVAILABLE

UNEQUAL CELL SIZE ANOVA AND THE MEANING OF THE INTERACTION

Ernest A. Rakow, The University of Memphis

Analysis of variance (ANOVA) is a frequently used statistical procedure in education and the social sciences. Very often the use of ANOVA involves situations with unequal cell sizes. When confronted with data to analyze from an unbalanced design, the researcher should very carefully select the method/option in the statistical package for estimation of the sums of squares.

As pointed out by Halpin, Carwile & Halpin (1991) the hypotheses being tested in unbalanced designs may not be precisely as expected. In fact, as will be shown, conflicting results may occur depending upon the options chosen in the analysis. The problem of analyzing non-orthogonal designs has been recognized as an issue since the 1950's when Scheffe published his book The Analysis of Variance (1959).

In this paper the cell means shown below will be used as an example to illustrate the points being made. Thus frequent reference will be made to the example in which independent variable **A** has two levels (two rows) and independent variable **B** has three levels (three columns)

Cell Means for ANOVA Example

	B 1	**B 2**	**B 3**	**Marginal**
A 1	51.0	41.5	48.5	47.0
A 2	59.0	54.5	45.5	53.0
Marginal	55.0	48.0	47.0	50.0

One approach presented in textbooks for doing ANOVA is such situations has been the use of unweighted means analysis. In this

approach each cell is given equal weight when calculating the marginal means and thus the sums of squares for main effects. In this case the marginal means in the example show are correct. The "n" used in calculating the sums of squares for each main effect and the interaction then is the harmonic mean of the cell frequencies. (Kennedy & Bush, 1985; Keppel, 1991; Keppel & Zedeck, 1989; Scheffe, 1959).

When we assume that all three null hypotheses are correct, the implication is that the grand mean, 50 in our example, can be substituted for all means in the table and that all variation from these means is random error. When the null hypotheses of equal means is rejected only for variable **A**, the implication is that all means in the first row have the same value, 47 in the example, the means in the second row all have the same value of 53, and the means in the marginal row are all the grand mean of 50. When the null hypotheses of equal means is rejected only for variable **B**, the implication is that all means in the first column have the same value, 55 in the example, the means in the second column have the same mean of 48, the means in the third row all have the same mean of 47, and the means in the marginal column are all the grand mean of 50. Rejection of the null hypotheses for both main effect of **A** and **B** (but not interaction) implies that the information from the grand mean and all marginal means is sufficient to explain variation in the data. In that case the implication is that the cell means in the example would be estimated as being 52, 45, 44 for the first row and 58, 51, and 50 for the second row. This is referred to as the additive model.

Interaction effects are defined as the difference between the actual cell means and those from the additive model. Thus, in the example, the interaction effects are -1.0, -3.5, and 4.5 for the

first row and are +1.0, +3.5, and -4.5 for the second row. These interaction effects are the unique combination effects of that combination of the two independent variables on the dependent variable. This is the non-additive model. As long as there are equal numbers of observations contributing to each cell mean these relationships hold and the estimation of sums of squares on the dependent variable associated with the three sources are all independent.

However, when the number of observations in the cells are not all the same, the sources of variation discussed above become correlated. The degree of effect on the analysis is determined by the degree of inequality of the cell n's. "In non-orthogonal designs, it is not possible to partition unambiguously the proportion of variance accounted for, or the regression sum of squares, into components attributable to each of the terms of the design" (Pedhazur & Schmelkin, 1991, p.536). Proponents of the unweighted means analysis argue that it's use is appropriate when the researcher can assume that in the population there should be equal numbers in each cell. The inequality in this study only occurs due to random factors which are not associated with any variable under consideration. The estimates which are independent can still be made. Otherwise a different approach should be followed. Pedhazur, while advocating a different approach, claims "all analytic solutions to the problem of unequal cell frequencies are based on the assumption that subject attrition is random" (Pedhazur & Schmelkin, 1991, p.536).

The difficulty created by unequal **n**'s can be illustrated in the example by using cell frequencies of 12, 10, and 8 for the first row and 5, 10, and 16 for the second row. When cell means

are weighted by the cell **n**'s, the marginal row means become 47.17 and 50.58, which clearly are not centered around the original grand mean of 50 (assuming equal cell **n**'s). The column marginal means are also changed, becoming 53.35, 48.00, and 46.50, with the new grand mean of 48.90. The interaction effects are still the variation in the cell means not explained by the grand mean and marginal means, but this is no longer easily estimated. Thus, what precisely are the hypotheses being tested when the **n**'s are unequal (Halpin, Carwile, & Halpin, 1991)?

Those authors most clearly explaining the alternatives for analyzing designs with unequal **n**'s do so by emphasizing comparisons of models via the use of sums of squared errors for competing models also tend to use regression methodology to explicate the various models for non-orthogonal designs. (Bogartz, 1994; Kirk, 1982; Maxwell & Delaney, 1990; Winer, Brown & Michels, 1991, and Woodward, Bonett, & Brecht, 1990). Several of their points will be made in the remainder of this paper via example. Others even suggest the use of exploratory data analysis to further understand the findings from ANOVA (Hoaglin, Mosteller, & Tukey, 1985 and 1991), although their approach is not illustrated here.

Table 1 shows the design layout for the examples used in the remainder of this paper. The same cell means are continued to be used in the five cases of various patterns of cell frequencies. For each case, the required number of scores were generated within each cell to produce the desired cell means and a mean square error of nearly 100.00. This table also shows the weighted marginal means associated with these **n**'s. Case 1 uses equal frequencies of 10 per cell. Notice that the more unequal the cell **n**'s the greater the impact on the weighted marginal means.

The data for each case was analyzed via SPSS using the ANOVA

procedure using in sequence the options for hierarchical, experimental, and unique. The same data was analyzed via SAS using the options for type I, type II, type III, and type IV. In all cases SAS type I and SPSS hierarchical produced the same results. SAS type II and SPSS experimental produced the same results. SAS type III, SAS type IV and SPSS unique produced the same results.

Table 2 presents the results for Case 1, the equal **n**'s situation. The different options all produced the same results. Table 3 presents the results for Case 2, the proportional **n**'s situation. Here the SAS type I and type II and SPSS experimental and hierarchical options produced the same results. The SAS type III and IV and SPSS unique solutions were equivalent and were different from the other analyses for this case only for source **A**.

Table 4 presents the results for Case 3, an unequal **n**'s situation. In this situation SAS type I and SPSS hierarchical options produced the same results. SAS type II and SPSS experimental options produced the same results, which for this case differed from SAS type I only with respect to source A. SAS type III and SPSS unique options produced the same results which agree with the other analyses only for the interaction effects. In this case the interaction is not significant even though the cell means are the same as the other cases.

Table 5 presents the results for Case 4, an unequal **n**'s situation. In this situation SAS type I and SPSS hierarchical options produced the same results. SAS type II and SPSS experimental options produced the same results, which for this case differed from SAS type I only with respect to source A. SAS type III and SPSS unique options produced the same results which agree with the other analyses only for the interaction. In this case also the interaction is not significant. In fact, the interaction

mean square is the same for Case 4 as it was in Case 3.

Table 6 presents the results for Case 5, an unequal **n**'s situation. In this situation SAS type I and SPSS hierarchical options produced the same results. SAS type II and SPSS experimental options produced the same results, which for this case differed from SAS type I only with respect to source A. Source A was not significant for Type I but was for Type II. SAS type III and SPSS unique options produced the same results which agree with the other analyses only for the interaction.

Why do these different options produce different solutions? The answer is in the approaches for dealing with the lack of independence. Table 7 shows the coded vectors which could be used to analyze these data via regression analyses (which were done, but are not shown in tables in this paper). Typically one type of coding would be used. For factorial ANOVA effect coding or orthogonal coding are preferred over dummy coding. Essentially each row in a coding matrix is repeated for each observation within a cell. When orthogonal codes are used with equal **n**'s the correlations among the vectors are all zero as shown in Table 8 for Case 1. For Case 2 (proportional **n**'s) the codes for main effect **A** correlate zero with those for main effect **B**; however there are some small correlations between the main effect codes and the interaction codes. This pattern of correlations is why SAS type I and type I solutions were the same and are different from SAS type III and IV.

For Case 3, Case 4, and Case 5 there are correlations among all of the vectors. The patterns of correlations lead to the various solutions shown above. Notice that for Case 3 and Case 4 the degree of relationship is the same for each corresponding pair of vectors but the signs of several correlations have changed.

This was caused by simply alternating the pattern of **n**'s within the rows between these cases. Also, this is why the ANOVA solution for Case 3 and Case 4 via SPSS unique was the same. For Case 5 the magnitude of correlation for the **A** and $\mathbf{B}_1$ of -.30 is the reason the ANOVA results for the **A** variable were so different between the SAS type I and type I solutions.

Conclusions

1. The various ANOVA options all lead to the same solution for interaction.

2. For SAS type I and SPSS hierarchical options the solution for the first main effect variable is the same as if that variable were the only one in the model (i.e. the same as regression in which the vectors for that source are entered first). This sum of squares and mean square would be the same as computed in a one factor ANOVA.

3. For SAS type I and SPSS hierarchical and for SAS type II and SPSS experimental options the sums of squares, means squares, and **F**'s are the same for the second factor. This solution can be thought of as the sum of squares explained by both main effects controlling for the sum of squares explained by the first factor (i.e. SS main effects minus SS(A)). For SAS type II and for SPSS experimental options the sum of squares for the first factor can be thought of as the sum of squares explained by both main effects controlling for the sum of squares explained by the second factor (i.e. SS main effects minus SS(B)).

4. SAS type III & IV and SPSS unique options extract sum of squares for the first factor which is the same as if the unweighted means analysis were done. Although several authors encourage this approach under appropriate conditions, this does not seem appropriate because the sums of squares for a main effect is

computed after controlling for the other main effect and the interaction. Given the definition of interaction above, it is not logical to enter a main effect in a sequence following the interaction. Although, if the interaction is significant, the primary focus should be on the cell means, ignoring the marginal means.

5. It appears that in most situations the SPSS experimental & SAS Type II would be preferred. Primary interest in specific hypotheses might justify the other approaches, but this author then recommends comparison of such results with those from the experimental approach.

REFERENCES

Bogartz, R. S. (1994). An introduction to the analysis of variance. Westport, CT: Praeger Publishers.

Hoaglin, D. C., Mosteller, F., & Tukey, J.W. (1991). Fundamentals of Exploratory Analysis of Variance. New York, NY: John Wiley and Sons.

Hoaglin, D. C., Mosteller, F., & Tukey, J. W. (1985). Exploring data tables, trends, and shapes. New York, NY: John Wiley and Sons.

Halpin, G., Carwile, R., & Halpin, G. (1991). So you have chosen an unequal cell size ANOVA option: Do you really know what you have? Lexington, KY: MSERA presentation.

Kennedy, J. J. & Bush, A. J. (1985). An introduction to the design and analysis of experiments in behavioral research. Landam, MD: University Press of America, Inc.

Keppel, G. (1991). Design and analysis: A researchers handbook. Englewood Cliffs, NJ: Prentice Hall.

Keppel, G. & Zedeck, S. (1989). Data analysis for research designs. New York, NY: W. H. Freeman and Company.

Kirk, R. E. (1982). Experimental design: procedures for the behavioral sciences. Pacific Grove, CA: Brooks/Cole Publishing Company.

Maxwell, S. E. & Delaney, H. D. (1990). Designing experiments and analyzing data. Pacific Grove, CA: Brooks/Cole Publishing Company.

Pedhazur, E. J. & Schmelkin, L. P. (1991). Measurement, design, and analysis: An integrated approach. Hillsdale, NJ: Lawrence Erlbaum Associates.

Scheffe, H. (1959). The analysis of variance. New York, NY: John Wiley & Sons, Inc.

Winer, B. J., Brown, D. R., & Michels, K. M. (1991). Statistical principles in experimental design. New York, NY: McGraw Hill, Inc.

Woodward, J. A., Bonett, D. G., & Brecht, M. (1990). Introduction to linear models and experimental design. San Diego, CA: Harcourt Brace Jovanovich, Inc.

Table 1. Cell Frequencies and Means and Marginal Means

Cells			CASE				
			1	2	3	4	5
A	B	Mean	n	n	n	n	
1	1	51.0	10	8	8	12	12
1	2	41.5	10	10	10	10	10
1	3	48.5	10	10	10	8	8
2	1	59.0	10	8	12	10	5
2	2	54.5	10	10	10	10	10
2	3	45.5	10	10	8	10	16

Weighted Marginal Means

A_1	47.0	46.7	46.7	47.2	47.2
A_2	53.0	52.6	53.9	52.6	50.6
B_1	55.0	55.0	55.8	54.2	53.3
B_2	48.0	48.0	48.0	48.0	48.0
B_3	47.0	47.0	47.2	46.8	46.5
Grand Mean	50.0	49.6	50.4	49.8	48.9

Table 2. ANOVA Results, Case 1

Equal Cell Frequencies: 10, 10, 10, 10, 10, 10

Source	Mean Sq	F	Pr>F
A	540.00	5.40	0.024
B	380.00	3.80	0.029
A*B	335.00	3.35	0.043
Error	100.00		

Same results for SAS (all 4 types) and SPSS (3 options)

Table 3. ANOVA Results, Case 2

Cell Frequencies: 8, 10, 10, 8, 10, 10

SAS Type I & Type II and SPSS Experimental and Hierarchical

Source	Mean Sq	F	Pr>F
A	480.29	4.80	0.033
B	326.43	3.26	0.047
A*B	332.86	3.33	0.044
Error	100.04		

SAS Type III & IV and SPSS Unique

Source	Mean Sq	F	Pr>F
A	498.46	4.98	0.030
B	326.43	3.26	0.047
A*B	332.86	3.33	0.044
Error	100.04		

Table 4. ANOVA Results, Case 3

Cell Frequencies: 8, 10, 10, 12, 10, 8

SAS Type I and SPSS Hierarchical

Source	Mean Sq	F	Pr>F
A	747.81	7.48	0.009
B	354.10	3.54	0.036
A*B	311.36	3.11	0.053
Error	100.03		

SAS Type II and SPSS Experimental

Source	Mean Sq	F	Pr>F
A	569.49	5.69	0.021
B	354.10	3.54	0.036
A*B	311.36	3.11	0.053
Error	100.03		

SAS Type III & IV and SPSS Unique

Source	Mean Sq	F	Pr>F
A	511.49	5.11	0.028
B	359.94	3.60	0.034
A*B	311.35	3.11	0.053
Error	100.03		

Table 5. ANOVA Results, Case 4

Cell Frequencies: 12, 10, 8, 8, 10, 10

SAS Type I and SPSS Hierarchical

Source	Mean Sq	F	Pr>F
A	423.06	4.23	0.045
B	378.40	3.78	0.029
A*B	311.36	3.11	0.053
Error	100.03		

SAS Type II and SPSS Experimental

Source	Mean Sq	F	Pr>F
A	569.49	5.69	0.021
B	378.40	3.78	0.029
A*B	311.36	3.11	0.053
Error	100.03		

SAS Type III & IV and SPSS Unique

Source	Mean Sq	F	Pr>F
A	511.58	5.11	0.028
B	359.94	3.60	0.034
A*B	311.35	3.11	0.053
Error	100.03		

Table 6. ANOVA Results, Case 5

Cell Frequencies: 12, 10, 8, 5, 10, 16

SAS Type I and SPSS Hierarchical

Source	Mean Sq	F	Pr>F
A	177.69	1.78	0.188
B	372.07	3.72	0.030
A*B	344.28	3.44	0.039
Error	100.05		

SAS Type II and SPSS Experimental

Source	Mean Sq	F	Pr>F
A	430.31	4.30	0.042
B	372.07	3.72	0.030
A*B	344.29	3.44	0.039
Error	100.05		

SAS Type III & IV and SPSS Unique

Source	Mean Sq	F	Pr>F
A	482.98	4.83	0.032
B	302.40	3.02	0.056
A*B	344.29	3.44	0.039
Error	100.05		

Table 7. Coded Vectors for doing ANOVA via Regression

			Dummy Codes				
A	B	$\bar{X}$	A	B_1	B_2	AB_1	AB_2
1	1	51.0	0	1	0	0	0
1	2	41.5	0	0	1	0	0
1	3	48.5	0	0	0	0	0
2	1	59.0	1	1	0	1	0
2	2	54.5	1	0	1	0	1
2	3	45.5	1	0	0	0	0

			Effect Codes				
A	B	$\bar{X}$	A	B_1	B_2	AB_1	AB_2
1	1	51.0	-1	1	0	-1	0
1	2	41.5	-1	0	1	0	-1
1	3	48.5	-1	-1	-1	1	1
2	1	59.0	1	1	0	1	0
2	2	54.5	1	0	1	0	1
2	3	45.5	1	-1	-1	-1	-1

			Orthogonal				
A	B	$\bar{X}$	A	B_1	B_2	AB_1	AB_2
1	1	51.0	-1	1	1	-1	-1
1	2	41.5	-1	0	-2	0	2
1	3	48.5	-1	-1	1	1	-1
2	1	59.0	1	1	1	1	1
2	2	54.5	1	0	-2	0	-2
2	3	45.5	1	-1	1	-1	1

Table 8. Correlations Among Orthogonal Coded Vectors

Case 1. Equal N's: 10, 10, 10, 10, 10, 10

	A	B_1	B_2	AxB_1	AxB_2
A	1.00				
B_1	.00	1.00			
B_2	.00	.00	1.00		
AxB_1	.00	.00	.00	1.00	
AxB_2	.00	.00	.00	.00	1.00

Case 2. Proportional N's: 8, 10, 10, 8, 10, 10

	A	B_1	B_2	AxB_1	AxB_2
A	1.00				
B_1	.00	1.00			
B_2	.00	-.06	1.00		
AxB_1	-.09	.00	.00	1.00	
AxB_2	-.05	.00	.00	-.06	1.00

Case 3. Unequal N's: 8, 10, 10, 12, 10, 8

	A	B_1	B_2	AxB_1	AxB_2
A	1.00				
B_1	.13	1.00			
B_2	.03	.03	1.00		
AxB_1	.04	.05	.09	1.00	
AxB_2	-.03	.09	.02	.03	1.00

Case 4. Unequal N's: 12, 10, 8, 8, 10, 10

	A	B_1	B_2	AxB_1	AxB_2
A	1.00				
B_1	-.13	1.00			
B_2	-.03	.03	1.00		
AxB_1	.04	-.05	-.09	1.00	
AxB_2	-.03	-.09	-.02	.03	1.00

Case 5 Unequal N's: 12, 10, 8, 8, 10, 10

	A	B_1	B_2	AxB_1	AxB_2
A	1.00				
B_1	-.30	1.00			
B_2	.01	-.10	1.00		
AxB_1	-.14	-.02	-.22	1.00	
AxB_2	.01	-.21	-.01	-.10	1.00